Foundation Proposal for a Chymical Encyclopedia, Database, and Repository (CEDR)

Grant Number: PW-228135-15

White Paper Report

James R. Voelkel and William R. Newman

with contributions by Michelle Dalmau and Wallace Hooper

Table of Contents

Introduction

Chemical terminology changed radically as part of the "Chemical Revolution" at the end of the 18th century. This linguistic upheaval demarcates two periods, a relatively recent modern-looking chemical literature and a much more mysterious historically contingent set of signifiers. The tradition of chymical practice, however, was continuous, stretching back for centuries and—for some operations—millennia (as the *Oxford English Dictionary* points out, the term "chymistry" has been adopted "to differentiate the early, transitional science from the discipline of 'modern' chemistry as practiced from the 18th cent. onward"). Historians and other humanists from a wide range of disciplines thus stand on the wrong side of a historical divide when it comes to understanding what materials and processes chymists and a variety of allied artisans were writing about and working with. In effect we are dealing with a dead language whose terms can only be understood fully by a combination of historical philology and replication of old chemical techniques and apparatus. We propose to bridge this divide with a project that uses the historical actor's own reference works and replication of their described processes to track the historic meaning of changing vocabulary and translate that language into modern terms. We anticipate that this work will be of interest to historians of science and technology—obviously—and also anyone interested in material history, whether it be conservators trying to replicate period pigments, instrument makers investigating the properties of historic varnishes, or historians of any kind trying to understand historic materials and their properties.

We ultimately aspire to create an integrated encyclopedia, database, and repository of recipe literature and dictionaries dealing with materials before the nineteenth century. The project itself has a tripartite organization: a repository containing a series of digitized sources, books and manuscripts dating from late antiquity to the turn of the 19th century, and videos and reports of modern replications and reconstructions of historic chymical experiments and operations; a database of chymical terms, referring to materials, apparatus, and processes, culled computationally from the contents of the repository and other available sources—as far as possible—and curated, interpreted, and extended by expert researchers; and an edited encyclopedia with definitive articles including videos of replications where possible to make the earlier sources intelligible. We plan to use a variety of original sources from different centuries as stepping stones, tracking the continuity in and change of terms over time. The original sources will provide the terms and headwords to populate the database, the database will inform the authors of the encyclopedia, and the on-line encyclopedia will disseminate the findings to audiences both academic and public.

To organize a broad coalition including historians, conservators, metallurgists, and historically-minded chemists, we spent the largest part of the Foundations grant organizing an international workshop at the Chemical Heritage Foundation to bring together stake-holders to advise us on the most useful organization and workflows for the encyclopedia. Building on their input, we planned to develop proto-types of the Encyclopedia with sample articles prepared by participants, of the Repository, with a small number of digitized original sources, and a representation of the Database with seed terms to build upon. It was our hope that this organizational workshop and proof-of-concept project would provide the nucleus of a working community and the foundation for a later implementation grant.

The primary purpose of the foundations grant was to hold a planning workshop at the Chemical Heritage Foundation to solicit advice on all aspects of the planned Chymical Encyclopedia Database and

Repository and to build a foundation of participants from a variety of fields. The workshop was held May 4-8, 2016. See Appendix A for the list of participants, and Appendix B for the program.

The workshop commenced with an introduction of the participants, with each giving a 10-15-minute presentation about themselves and their work. Since the participants were drawn from such a wide range of disciplines—ranging from history of alchemy and classics, to art conservation, historical archeology, and digital humanities—we felt that time needed to be committed to learning about one another and our various interests in and approaches to historical "chymistry". After introductions, the workshop had presentations and discussions on different aspects of the project covering the encyclopedia model, the repository, the database, historical replication, chemist-historian collaboration, and the vision of the CEDR project.

The Encyclopedia Model

From the outset, one model for CEDR has been the highly-successful online Stanford Encyclopedia of Philosophy (SEP), and the workshop was very fortunate to secure the participation of Uri Nodelman, a senior editor at SEP to discuss the SEP model and comment on the similarities and differences in the aim and content of SEP and the proposed CEDR. He framed his presentation with the mission statement of SEP:

> The Stanford Encyclopedia of Philosophy organizes scholars from around the world in philosophy and related disciplines to create and maintain an up-to-date reference work.

He used this statement as a springboard to highlight the importance of *people*. His work at SEP, he said, is less as an editor and more as a community organizer. The importance of *a community of people working together* is central to the success of the project. Having seen everyone's introductory presentations the previous day, he allowed that from what he could see we have a healthy nucleus of such a community.

It is important that the motives and audience for this community match the goals of the project. Nodelman said that SEP contributors worked to produce a reference work, "for themselves, colleagues, students, and the general public in that order." And he said that the level of treatment was aimed at the convenient fiction of the "advanced undergraduate," in that the goal of SEP entries is to provide an introduction to a subject and its current bibliography.

In this, he began to contrast what CEDR aspired to versus the SEP. Though there are exceptions, he characterized SEP entries as *tertiary sources* in that they are not meant to present original research but summarize the current state of the secondary literature. From what he had seen of the CEDR proposal, it seemed to him that the goal of CEDR was sufficiently geared toward original research to make it more of a secondary source, and because of this, he asked the question whether it was really accurate to characterize CEDR as an "encyclopedia."

Dr. Nodelman gave an in-depth presentation of the editorial structure and workflow at the Stanford Encyclopedia of Philosophy. While the details of the advisory committee, editorial board, editorial procedures and workflow were more suited to a large, mature project like the SEP, there were nonetheless many useful details that informed us about how CEDR would likely need to be run.

Among these was that, in an important sense, *the work on an entry is never done*. Even after a topic has been identified, an author chosen, the article written and passed through multiple rounds of reviews, and made public, the end of that process only starts the clock on a 4-year deadline to revise the article, checking the current state of the literature and updating the entry as necessary. This brought Dr. Nodelman back to the importance of people to the project. When the author accepts the invitation to write the article, she or he also accepts an ongoing commitment to maintain the entry "until death." Though this was said in jest, his detailed description of succession-planning for entries made clear that the commitment to maintain an entry was important and long lasting.

The role of the editorial board was also something CEDR needs to consider. The editorial board at SEP determines the shape and direction of the encyclopedia, determining what new topics to solicit entries for, which entries to consolidate, and which to retire to the archive. The editorial board is also an important line of defense: since every entry is commissioned, no unsolicited entries are accepted, and the authors of entries are chosen by the editors. Such a stance would be a helpful bulwark against "contributors" with fringe alchemical beliefs. On the other hand, the editorial process of the SEP highlights how much larger a community philosophy is than chymistry. There were barely enough attendees at the workshop to make up an editorial board, let alone commission entries for every topic deemed worthy of being included in CEDR.

While Dr. Nodelman's comments about secondary versus tertiary sources are well taken, it is not the case that print resources along the lines of what CEDR envisions have not been partially implemented in the past. The most recent example is Claus Priesner and Karin Figala's 1998 *Alchemie: Lexikon einer hermetischen Wissenschaft*; in addition, a series of German monographs published by the Pharmaziegeschichtliches Seminar in Braunschweig in the 1960s provided a wide variety of entries on chemical and medicinal materials employed in the pre-modern era. Indeed, one can find precedents to the encyclopedic goals of CEDR as far back as Charles Singer's 1954 *History of Technology* if not before. Unlike the *SEP*, however, CEDR will have to draw its content from specialists in a wide range of disciplines including history of science, experimental archeology, art history, chemistry, and other fields. Although much of the research has already been done, it cannot simply be harvested from a single academic discipline like philosophy. By its very nature, CEDR is interdisciplinary.

Repository

The vision for the R in CEDR is a repository that will source the encyclopedia and database components of the end-user, online encyclopedia. The repository would be mined by contributors for both primary and secondary resources to aid in the composition of encyclopedia articles, and by the back-end, infrastructure for the extraction of headwords that would populate the database thereby creating article stubs and future access points for readers of the online encyclopedia. As the coverage for the encyclopedia could cover many centuries of chemical literature predating the 18th century, the creation of the repository needs to confront several challenges: diverse formats of resources from manuscripts to early printed books, diverse genres from dictionaries to scientific journaling, and many languages, including non-Latin script. In addition to challenges related to form, genre, and language, some of the resources could be copyright protected. Michelle Dalmau, Co-Director for the Indiana University (IU) Institute for Digital Arts & Humanities, and Head of Digital Collections Services for the IU Libraries, presented a paper entitled "Text Repositories for Digital Humanities Projects" at the CEDR workshop.

The paper defined ‘digital repositories’ for the scholarly audience in lay terms with an emphasis on data curation, re-use, and discovery. The paper went on to illustrate existing repository services provided by IU Libraries for the curation, preservation, and access to media, images, and texts – all components of the multimedia CEDR project. Finally, a short introduction on text encoding was provided, with an exploration of pros and cons as a segue to exploring potential repository frameworks for the CEDR project: MONK (Metadata Offer New Knowledge, Mellon funded project), TAPAS (TEI Archiving, Publishing and Access Services, funded by NEH and IMLS), TextGrid (funded by the German Federal Ministry of Education and Research), and the Hathi Trust Digital Library / Hathi Trust Research Center (funded by NEH and Mellon). Gabriele Ferrario, Research Associate, Taylor-Schechter Genizah Research Unit, Cambridge University Library, led the presentation on multilingualism based on his contributions to text mining for the Cairo Genizah project and issues to consider when dealing with different alphabets and reading orientation. The presentation ended with the following discussion prompt:

> Projects of this kind must all confront a central set of strategic concerns and design challenges, including questions about how much uniformity to impose upon the data, how to accommodate variation, how to create interoperability layers and tools that can operate meaningfully across multiple data sets (Blanke et al. 2011), and how to manage issues of sustainability (of both the data and the service itself).[1]

The resulting discussion about the repository function for CEDR raised several issues that ultimately hinged on understanding the audience for the repository. Will the repository be accessible to readers of the encyclopedia as a way to trace sources for the encyclopedia article or search for concepts that may manifest in various ways (i.e., sericon/red lead)? Or will be the repository serve as a resource for contributors only, a work-space of sorts, for scholars and experts to work through the evidence? Naturally, discussions of copyright ensued, especially as the publication of critical editions that include translations of primary sources, commentary and analysis, and related apparatus is a common form of scholarly output for historians of science. To provide an entirely open repository, could only parts of the text be provided (i.e., primary source reproductions) or could texts be searchable though not displayed as full-text for the reader? Though originally proposed as a text-only repository, discussions ensued about prospects for a multimedia repository that would include primary and secondary sources and ways to distinguish not only format but also type of source materials. Finally, discussions around the importance of extending the repository to include non-Latin-based alphabets could not be overstated, but the foreseen and imagined technical challenges indicate that further exploration of this topic is necessary. The workshop participants emphasized that one of the main goals for mining texts from across place and time is to locate the usage of a chemical/material term of concept whose nomenclature changes over time, which is further compounded by the use of metaphors or obscure synonyms. As solutions and approaches were considered with this one goal in mind, two options presented themselves: 1) concordance and 2) an ontology or thesaurus of terms and concepts that would include normalization of terms across time and languages. The former could be an interim solution that relies of big data text mining methods, likely by language, to serve as an index to the texts as a whole. The latter would rely on a combination of automatic and human processing of the texts, likely through text encoding or tagging, that could take a multi-lingual approach.

---

[1] Julia Flanders and Scott Hamlin, “TAPAS: Building a TEI Publishing and Repository Service,” Journal of the Text Encoding Initiative [Online], Issue 5 | June 2013. URL :http://jtei.revues.org/788 ; DOI : 10.4000/jtei.788

The consensus was that at a minimum the repository needed to provide access controls and support formats beyond texts, especially as encyclopedia contributors will be providing photographs and videos of recreated experiments. Based on the feedback from the workshop participants and our limited resources for the prototype, an integrated, seamless repository solution with advanced text-based processing for CEDR will require a significant investment in multilingual domain and technical expertise. For the prototype, the CEDR project team will rely on three existing repository services provided by the IU Libraries that each feed directly into our Fedora digital object repository: Pages Online (for the texts, http://pages.dlib.indiana.edu), Media Collections Online (for the video reenactments of experiments, http://media.dlib.indiana.edu), and Image Collections Online (for photographs or other images that would accompany encyclopedia entries, http://www.dlib.indiana.edu/collections/images/) with an initial focus on Pages Online, which the CEDR project has launched with volume one of the multi-volume *Dictionnaire de chymie* of Pierre-Joseph Macquer (1766-1789): https://webapp1.dlib.indiana.edu/cedr. The current version of Pages offers access controls so that restricted access ranging from a designated group to no access at all is possible at the item-level. Because Pages is built on the open-source Fedora/Hydra technology stack that is increasingly being adopted by academic librarians and other cultural heritage sectors, the possibility to extend the functionality for Pages to generate concordances, integrate a thesaurus, and connect to multiformat resources for a seamless discovery interface.

Database

In the conceptual space between the public encyclopedia and the repository, we envisioned the construction of a large database of materials and recipes that will allow contributors and participants to track and compare processes and artisanal operations as they were developed over long historical time. To give one example of how this might work, consider the fact that alongside natural azure or ultramarine produced from lapis lazuli since late antiquity, medieval practitioners and writers also described another “azure,” a synthetic form made from silver subjected to acidic vapors. The azure-like tint of this silver-blue was due to copper impurities that it contained, and had no chemical similarity to real azure. The database fields would record both actors’ categories and modern chemical information, however, and would therefore return queries about “azure” with both the genuine material and its ersatz imitations (noting the difference between them of course). As the database grows over time, more sources will be entered and the chronological coverage will increase, making it possible in many cases to arrive at the earliest testimonies for products such as the synthetic azure or silver-blue. A database of materials and recipes of this kind would therefore provide a crucial resource for the encyclopedia authors, since it would allow them to chart the evolution of processes over their history.

Our initial ideas of the CEDR database were informed by the design and experiences of the Cologne database for painting materials and techniques, the *Kunsttechnologische Rezepte des Mittelalters und der frühen Neuzeit*, which is the groundbreaking work of Dr. Doris Oltrogge of the Cologne Institute of Conservation Sciences.[2] We were fortunate that Dr. Oltrogge could attend the Philadelphia workshop and participate in our discussions. She reviewed the design and purposes of the Cologne database for us in her introductory remarks.

---

[2] Doris Oltrogge, “The Cologne database for painting materials and reconstructions,” in *Art of the Past: Sources and Reconstructions*, eds. Mark Clark, Joyce H. Townsend, and Ad Stijnman (London: Archetype Publications, 2005), 9-15.

In the face of substantial amounts of published and extant manuscript materials on painters' recipes in Germany and Europe, much of which has not been edited, Dr. Oltrogge's group designed the Cologne database as a vehicle to extract and collect the recipes themselves, and make them available for computer-based searches and reports. The recipes have been transcribed directly into fields in the database, with references to their sources. To address problems of reference and language like the confusions involved in the synthetic azure example above, Oltrogge's group constructed a thesaurus in the database to inform user queries. Oltrogge's solutions for search and organization of the recipe materials are obviously applicable to CEDR's problems. However, given that we see considerable value in developing and drawing upon an extensive repository and on available external sources, we are more likely in the CEDR database to depend on references, cross-references, and external links, and less on actual internal transcriptions of recipe materials except in occasional cases. Nonetheless, just like the Cologne database, one of the core purposes of the CEDR databases must be to support researchers in gaining organized access to a large store of recipe materials and historic commentary, in sorting those materials appropriately, and conducting fruitful search operations over the whole store.

The decision to extract and transcribe just the painters' recipes is one response to the enormous amount of relevant historical literature. Another kind of approach, already being explored by the Chymistry of Isaac Newton Project at Indiana University, and by the Taylor-Schechter Genizah Research Unit at Cambridge University, is to apply computational methods for large corpora analysis. We anticipate that the application of computational methods will be an important research strategy for the future of the CEDR project.

Prof. William Newman and Dr. Wally Hooper at the Chymistry of Isaac Newton Project, for example, used latent semantic analysis in 2011-12 to assemble a concordance of passages from across the Newton alchemical corpus that were related to the headwords of Newton's own chymical glossary, and his *Index Chemicus*, where the related passages are collated in a ranked order using latent semantic analysis. However, because we have little additional information about the processes Newton was using and imagining, the resulting concordance is a bare alphabetical framework of headwords with their own ordered collections of collated fragments with citations. Each headword and its collection will no doubt require careful research before we begin to comprehend the meanings of the headwords for Newton himself from their contexts. Nevertheless, each collection is still a trove of starting points, and, more importantly, each was easily collected by computation—a recognition that counts as one of several basic foundations of the CEDR proposal.

The Taylor-Schechter Genizah Research Unit uses innovative computational methods to work with a collection of more than two hundred thousand fragmentary manuscripts from Cairo and several other sources, which date mainly from the eleventh through nineteenth centuries with some older materials. To make the manuscripts more accessible, especially online, approximately ten thousand have been photographed. The unit's first computational project used the images to assist with document identification. Alert editorial inspection had uncovered many cases where a set of fragments could be identified as parts of a single, original document, based on shared orthographic, linguistic, and physical features. A basic question is how to identify other such separated documents without relying on the memories of the editors alone. The Genizah Unit now processes visible textual and physical features in the manuscript images to try to identify comparable fragments across a very large catalog of images with computational methods and are having success.

Dr. Hooper of the Chymistry of Isaac Newton Project and Dr. Gabriele Ferrario of the Genizah Research Unit both spoke on the problems of working with databases in historical, textual, and physical research work and the issues of updating, editing, and using databases to extract results.

Dr. Hooper talked about the basic organization of dictionaries and databases, drawing on lengthy professional experience assisting linguists with the design and implementation of field databases for Native American languages while a research scientist at the American Indian Studies Research Institute, and more recent experience with computational projects for the Chymistry of Isaac Newton Project. He argued that, because of the core linguistic and semantic nature of many of CEDR's interpretive problems, it would be most effective to design the CEDR database along the lines of a dictionary database with plenty of support for languages and symbols, for image and video analysis, and useful cross-referencing and external links. He discussed the various capabilities of relational databases for combined lexicographical, historical, archaeological, and chemical analysis, and described possible features that would be needed in the CEDR database (see Appendix D: Possible Model for a CEDR Relational Database). The CEDR database should be a flexible structure that: (1) can accommodate or provide access to harvested computational results as well as to the contents of the repository, and materials beyond, all of which can accumulate over time; (2) provides centralized support materials like tables of names and authorities, sort orders for all of the languages involved, and other useful tables like a thesaurus and table of synonyms to support search and organization; and, (3) is extensive and robust enough to support the simultaneous work of a large team of collaborators, and allow them to organize and analyze particular entries most appropriately. As long as the CEDR database tables, fields, and indexes are well designed, we will be free to choose from a wide range of database platforms for its eventual implementation.

Dr. Ferrario discussed problems that arise when a small editorial team attempts to catalog and index a very large collection of significant fragmentary manuscripts, and described the use of computational tools to address some of these problems, beyond the document identifications mentioned earlier. Each manuscript is assigned an identifier called a class mark and a catalog of class marks exists but the editors recognized that listings for most of the material were bare and unhelpful for users. They decided not to release a manuscript until a thorough description was available, including a transcription and possible citations from the secondary literature. An editor can produce around a thousand such descriptions a year, so the team has decades of work ahead of it at that rate. However, the Genizah materials are so important for several fields that there is now a sizeable literature which makes references to Unit collections. The team decided to augment their catalog and indexes by adding the citations to the respective class marks. They conducted a meta-analysis to construct a search corpus of secondary materials (all in English to begin with) and identified around 3,500 candidate books and articles. After securing rights, and training their algorithm to recognize Taylor-Schechter class marks in print, it segmented the search corpus into sections and pages, and located 6,322 mentions of T-S class marks in roughly 6,500 OCR pages. Once a class mark was located, the algorithm extracted the section from the secondary literature and linked it or associated it with that class mark, and a link would be added to the catalog, greatly increasing the value of those entries to users. The team went further and conducted a topic-modelling analysis of the associated sections from the search corpus, with the idea of adding topics to class marks where associations could be demonstrated. The analysis was designed to produce seventy-five clusters of co-occurring terms that would be identified as topics by the Bayesian process. Dr. Ferrario reported that while some clusters were easy to label and add to the catalog, many clusters made no sense as a topic and, so, could not be used. The topic-modelling experiment had mixed results, but the range of the Genizah Unit's computational work shows that computational methods can be successfully employed by small teams to elicit useful clues and information from large bodies of relevant material, and their results can be readily ingested into existing databases.

Historical Replication

Replication of historical processes to produce different materials in order better to understand the practice of chymistry is a major principle of the CEDR project. We thus had two presentations on replication from two slightly different perspectives. Lawrence Principe, a chemist and historian of science, gave a presentation on the production of white lead and other substances, and Marcos Martinón-Torres in conjunction with Nicholas Thomas gave a presentation on replication in chemical archaeology. The major difference between these two areas, as Prof. Martinón-Torres noted, is that historians of chemistry usually begin with a recipe they are trying to follow to produce a substance whose properties they might be uncertain of, whereas archeologists begin with an artifact with fixed properties that they try to reverse-engineer in order to understand the process by which it was produced, usually without any recipe. Apart from that fundamental difference in approach, many of the issues they discussed were similar, and we will try to treat them simultaneously, while leaving out many of the fascinating details of the chemistry of the Bologna stone or the source of chromium in artisanally-produced high strength steel.

Prof. Principe began with the problematic nature of texts. These were often (usually?) by scholars trying to record the actions of practitioners, and thus from the beginning there is a translation problem, both in terms of (1) technical terminology that may be misunderstood by the writer (e.g. the distinction between the technical and casual uses of 'melt' and 'dissolve') and (2) tacit knowledge of the practitioner that is not amenable to being written down. Readers of texts, on the other hand, can have knowledge insufficient enough that they simply demand more detail than is practical to record, or, especially in the case of modern chemists, they can make unwarranted assumptions that distort or override the actual content of the text. Indeed, he said that the facile application of modern chemical knowledge to an armchair reading of an historical text is "almost invariably wrong."

For these reasons, among others, there is a dynamic relationship between text and practice in historical replication. The text informs the replication, but at the same time results from the replication (or its failure) also informs how the text must be read. Indeed, in one of his most interesting examples, Prof. Principe discussed the Greek term ἀφιθοῦσιν, whose use is recorded only in Theophrastus. From the experience of replicating the production of white lead according to Theophrastus's account, he concluded that the meaning of the term is best understood as the technique technically known as 'levigation', which in turn informs the translation of this rare term from classical Greek.

Profs. Principe and Martinón-Torres concurred that a replication is an interpretation that is undertaken for explanation's sake. Prof. Principe outlined how a replication begins by stripping away incidental details to arrive at the core of the issue at hand. This painstaking work is iterative, and often involves adding back details thought to be irrelevant. Similarly, Prof. Martinón-Torres described experiments systematically changing one variable at a time in an attempt to isolate the key relevant element and achieve some predictive power.

Both speakers gave compelling presentations that highlighted the explanatory potential of replication of chemical procedures. Both emphasized replication's role as another historical tool. And both emphasized that—like other historical explanations—the results of replication are provisional. However, it was clear from both speakers' presentations that the work of replication was exacting and time consuming, which lead Prof. Smith to observe during the discussion that due to the very high level of effort required, it might be that only exemplary replications will ever make their way into CEDR.

Chemist-Historian Collaboration

One of the hopes of the CEDR project is to facilitate the collaboration of historians and chemists in investigating the historical use and development of chemical substances and processes. Seth C. Rasmussen, a professor of materials science at North Dakota State University, anchored our discussion of Chemist-Historian collaboration. Prof. Rasmussen was well situated to lead this discussion, as he is both a grant-winning, lab-leading research chemist and a chemist-historian, who has published numerous papers on historical materials.

Prof. Rasmussen began by drawing a general distinction between two types of chemists interested in history at all:

- Chemists with a strong interest in history, who introduce history into their lectures and publications, but who themselves recognize that they do not aspire to original research.
- Chemist-historians—far fewer in number—like himself, who have progressed sufficiently from the former state to be comfortable conducting and publishing historical research.

He then discussed briefly the hurdles of moving chemists up to the level of the chemist-historians suitable for making contributions to CEDR, of which there were dauntingly many. To begin with, simply, chemists are generally less interested in the history of their discipline than other scientists like, say, astronomers. Second, those who are interested in history tend to be interested in more recent, more intelligible to them history than the centuries envisioned for the CEDR project.

But the most daunting challenges are structural and institutional. Chemists who run research labs are very busy. They have an expected output of papers per year, and grants (with indirect costs that disproportionally support their universities) to get and maintain, students to supervise, and do not have much or any time left over to indulge in an interest in history. To make matters worse, chemists' historical publications are not seen by their departments or peers as publications that count toward their expected output. And finally, chemists without active grant-getting, paper-producing labs, quickly lose their labs, depriving them of the facilities which might be used in historical replication for CEDR.

Which brought him to the question of where to find possible chemist collaborators. He had the following suggestions:

- Existing chemist-historians, whom we could connect to through their existing groups, e.g. the HIST division of the American Chemical Society, the Historical Group of the Royal Society of Chemistry, or--as was pointed out in the discussion—the Society for the History of Alchemy and Chemistry (SHAC).
- Chemists with a strong interest in chemistry and a strong educational focus, such as might be found in undergraduate liberal arts programs.

- Chemists in related fields, e.g. chemical archeologists, or—as came up in discussion—chemists involved in art conservation.

The second of these groups was then the focus of a discussion of how to encourage and facilitate their involvement.

Chemists whose commitment is to education rather than research, such as those at undergraduate institutions, are relieved of the burdens of running a big research lab, but they still do research and have labs in order to educate undergraduates. The key to enlisting their support for CEDR, according to Prof. Rasmussen, would be to involve them in a project that would contribute to their careers, for instance, by making the replication of a historical procedure an educational activity. (Prof. Newman in the discussion pointed out that just such educational materials had been produced as part of the Chymistry of Isaac Newton project, at http://www.chymistry.org/chemlab/chemlab.pdf.)

Involving these researchers would require mentoring, both in pointing them toward a specific project and in teaching them the methods of historical research. To lower the barrier to entry, Prof. Rasmussen suggested having more experienced chemist-historians mentor younger chemists, or arranging training, either in the form of an online guide, video tutorials or a workshop. A representative of the Chemical Heritage Foundation proposed that CHF might be able to act as a mediator between its constituents—who demonstrate their interest in history through their affiliation with CHF—and historians involved with CEDR. At this point Prof. Principe pointed out that he had proposed just such a workshop to CHF without success. Perhaps in conjunction with the CEDR project this proposal for a workshop should be re-examined.

The CEDR Prototypes

Though the vast majority of the funds awarded were devoted to the highly successful CEDR Workshop, an important aspiration of the project was a proof-of-concept website that would implement the different elements of the Chymical Encyclopedia, Database, and Repository.

To provide a working web-based starting point for the prototype elements, we requested a group-project website on Indiana University's general Webserver (http://www.iu.edu/~cedr/cedr/), and created a home page and pages for People, Encyclopedia, Database, and Repository. (See Appendix C, Figures 1 and 2.) The IU group website provides a protected development environment and a production site, and generous storage space for streaming media. We also installed an instance of MediaWiki on the site so the project could have an informal wiki space for collaboration and preparations. After considering off-the-shelf options for a platform for the encyclopedia component, we decided to use an instance of Omeka, which was well supported and documented, and would allow us to consider the editorial and organizational issues that the CEDR initiative must address. The repository component presents a serious challenge because of the range of formats and media that interest us, but we had decided to scan Macquer's *Dictionary of Chemistry* and produce page images and OCR text as a reasonable example of what a CEDR repository would need to do. IU Digital Collection Services was able to arrange permission to ingest our page images and OCR text in the IU Pages Repository, which is designed to provide precisely that service, though it cannot handle media or other kinds of materials.

For the **Encyclopedia**, our goal was to showcase the types of encyclopedia entries we imagined would be the end result and public-facing part of the project. In order to write these entries, the contributors would have to already have conducted the necessary research, because there would not be the possibility of using the repository and database during the construction. We were happy to receive commitments during the workshop for a sufficient number of articles covering a gratifyingly wide range of types of subjects.

Matteo Martelli, the editor of the magnificent modern edition of *The four books of Pseudo-Democritus* (Leeds : Maney Publishing, 2013), pledged and delivered an entry on Pseudo-Democritus, which simultaneously fulfilled our desire for an entry in the *biographical* category and one on *alchemy in antiquity*. (See Appendix C, figures 3 and 4.)

William Newman, General Editor of the Chymistry of Isaac Newton Project and author of a forthcoming book on Newton's alchemy, pledged an encyclopedia entry on Newton's alchemy, which fulfilled our desire for an entry on an *individual's practice* or a *school of alchemy*. It is in preparation, and we anticipate that it will be published on the website within a few months.

In the meantime, Prof. Newman's student Megan Allen produced an entry on different metallic reguli of antimony produced by replicating the methods in Newton's alchemical notebooks. This entry satisfies the category *replication* and *substance*. (See Appendix C, figures 5 and 6.)

Marcos Martinón-Torres and Nicolas Thomas pledged an entry on cupels/crucibles. This will supply the perspective of experimental archeology on a topic in the category or *apparatus and technique*, as well as *artifact*.

Marlolijn Bol pledged an entry on artificial gemstones, a subject with which she has considerable experience. This topic, which is another topic in the category *replication* and *substance*, is of considerable interest because the making of artificial gemstones was a distinct subspecialty of alchemy from antiquity through the 18th century. Unfortunately, due to a change of institutional affiliation, the entry was delayed at least until the fall of 2017.

The **Repository** section of the prototype website was to contain primarily Macquer's 5 volume final edition of the *Dictionnaire de chymie* (1789). Shortly after the workshop, in light of our discussion of the complexity of implementing multi-lingual headwords in the database and the added difficulty of encoding a French language book as our first full work, the decision was made to switch to the most complete English edition of Macquer's dictionary, the 3 volume *A Dictionary of Chemistry, Containing the Theory And Practice of that Science* (London, 1777). The work was digitized at very high quality using the Chemical Heritage Foundation's 80MP Phase One camera, and the page images were uploaded to the IU Pages repository. (See Appendix C, figures 7, 8, 9, and 10.) Encoding of the first volume of the work was begun with OCR performed by ABBYY Finereader 14. Unfortunately, despite extensive training of the OCR engine with the 18th century font, it was not able to overcome the extreme similarity of long s's (ſ) and f's, and the project of proofing the OCR proved too labor-intensive. The transcript of the first volume has nonetheless been posted on the CEDR prototype website. The subsequent insertion of TEI markup, from which the headwords were meant to be automatically extracted, was therefore also not possible. Even during the workshop, however, there was active discussion about whether TEI encoding was too high a bar to expect for items in the repository (see Repository section above). Now that the page images are in the Pages system, another opportunity for OCR presents itself, though it is doubtful that another engine will be more successful with the persistent ſ/f confusion.

The second seed for the text repository was the TEI-encoded transcript of Isaac Newton's manuscript, Don b. 15. At the time of this white paper, the full transcript can be viewed on the Chymistry of Isaac Newton website, and a list of the headwords—nominally awaiting passages harvested from the repository or elsewhere—has been mounted on the CEDR website's Database page.

The final seed for the text repository was a partial transcript of the Othmer Library manuscript *Secreti naturali*. A transcript of the index of recipes (roughly equivalent to the headwords of a dictionary) was prepared by workshop participant Joel Klein. A good first run through, it awaits further editing. Toward that end, high quality tiffs of the relevant pages of the *Secreti naturali* have been uploaded to the CEDR Pages instance. It was not deemed necessary to upload the entire manuscript to the CEDR prototype, since the manuscript is already publically available in its entirety as part of the Bibliotheca Philadelphiensis at Openn.

In light of some technical, but more importantly conceptual issues, the prototype **database** did not develop as we had proposed. We had originally conceived that construction of the database would be an almost automated procedure, drawing headwords from the *Dictionary of Chemistry* on the basis of the TEI tags, and combining them with the headwords from Don b. 15 and the *Secreti naturali*. The workshop discussion concluded that the database would need to be much more thoughtfully curated (see discussion in Database section above). Wally Hooper presented a model of a possible relational database for CEDR and an ingestion procedure for computationally harvested material to stimulate discussion (see a summary of that model in Appendix D: Possible Model of a CEDR Relational Database), but the development of the prototype database remains unrealized at the time of writing.

Conclusion and Next Steps

Uri Nodelman asserted in his presentation that the key to a successful online encyclopedia was first and foremost a community of scholars working together. In that context, the camaraderie and unity of purpose exhibited by our diverse, international group of participants was the surest sign of the potential of the CEDR project. Getting stakeholders from history of chemistry, art conservation, experimental archaeology, chemistry, and digital humanities together *before* embarking on a large online project allowed the organizers to solicit important input about the structure and goals of the project, to have frank, open-ended discussions about problematic elements of the project from how to structure the text repository and database to the role of multimedia to challenges inherent in replication. Most importantly, it set the stage to get essential buy-in and participation, as evidenced in the gratifying response to our call for volunteers to contribute to the prototype site.

Our next steps can be divided into two categories: those that we can immediately begin implementing, and those that will require the additional funding of an implementation grant. The following goals may be grouped within the category of those that we can begin to implement immediately:

- Extend the content of the prototypes based on work by current participants. There are a number of encyclopedia entries that can be developed with fairly minimal labor, based on work already done in the form of scholarly articles by participants.

- Define a more formal editorial process that would facilitate recruitment of new participants. Part of the success of the *SEP* rests on the formal editorial structure of the project. Extensive vetting of entries by recognized experts in the field grants the same sort of credit that one receives from publishing in a top-notch scholarly journal. A similar structure needs to be put in place for CEDR, beginning with the creation of an editorial board.
- Determine the precise form that the database will take. For example, will it incorporate substantial bodies of text as the Cologne art history database does, or will that role be relegated to the repository? Also, since we envision the database as a working tool requiring expert knowledge, will it need its own editorial board?
- Thanks to a very recent grant from Indiana University (over $80,000), the CEDR team is now in the position of being able to acquire an X-ray Fluorescence Spectrometer which will allow us to perform elemental analyses of historical materials produced in a Chemistry Laboratory at IU. This will greatly facilitate the work of replication, since it will allow the team not only to reproduce the products of old recipes, but also to identify their composition. Such replication is actively being carried out currently by Newman and his graduate students.

Additional steps that will require further funding, presumably from an NEH implementations grant, include the following:

- Possible hiring of data-entry personnel to carry out markup of the texts in the repository. It remains to be determined whether the texts in the repository will receive full TEI markup or not: this will depend in part on the platform that we ultimately choose for the repository.
- The multimedia side of CEDR will require considerable work going forward, which cannot be supported entirely by internal resources.
- Similarly, the infrastructure of the site—much of which is available through Indiana University – may need to be supplemented by outside funding.
- The biggest foreseeable expense will be the hiring of a dedicated editor for CEDR. We base this expectation on the model of the *SEP*, which has employed a full-time editor virtually since its inception.

Appendix A: Participants at the CEDR Workshop, Chemical Heritage Foundation, May 4-8, 2016

1. Donna Bilak, Postdoctoral Scholar, Making and Knowing Project, Columbia University.
2. Marjolijn Bol, Postdoctoral Researcher, Department of Conservation & Restoration, University of Amsterdam.
3. Jenny Boulboullé, Postdoctoral Scholar, Making and Knowing Project, Columbia University.
4. Will Cowan, Head of Software Development (Library Technologies), Indiana University Bloomington Libraries.
5. Michelle Dalmau, Head, Digital Collections Services, Indiana University, Bloomington.
6. Michelle DiMeo, Curator of Digital Collections, Othmer Library of Chemical History, Chemical Heritage Foundation; contributor to the Recipes Project and co-editor of Reading and Writing Recipe Books 1550-1800.
7. Gabriele Ferrario, Research Associate, Taylor-Schechter Genizah Research Unit, Cambridge University Library.
8. Wally Hooper, Project Manager/Programmer/Analyst, Chymistry of Isaac Newton Project, Indiana University.
9. Joel Klein, Postdoctoral Scholar, Making and Knowing Project, Columbia University.
10. Matteo Martelli, Research Associate, Institut für Klassische Philologie, Humboldt-Universität zu Berlin.
11. Marcos Martinón-Torres, Professor of Archaeological Science, University College London.
12. William Newman, Professor of History and Philosophy of Science and Director of Chymistry of Isaac Newton Project, Indiana University.
13. Uri Nodelman, Senior editor, Stanford Encyclopedia of Philosophy.
14. Doris Oltrogge, Cologne Institute of Conservation Sciences, Cologne, Germany.
15. Nicola Pohl, Professor of Chemistry, Indiana University.
16. Lawrence Principe, Professor of History of Science and Chemistry, The Johns Hopkins University.
17. Peter J. Ramberg, Professor of History of Science, Truman State University.
18. Jennifer Rampling, Assistant Professor of History, Princeton University.
19. Seth Rasmussen, Professor of Chemistry, North Dakota State University.
20. Pamela Smith, Professor of History and Director of the Making and Knowing Project, Columbia University.
21. Nicolas Thomas, Faculty Member, Institut National de Recherches Archéologiques Préventives, Centre-île de france.
22. James R. Voelkel, Curator of Rare Books, Othmer Library of Chemical History, and Resident Scholar, Beckman Center for the History of Chemistry, Chemical Heritage Foundation.
23. Arie Wallert, Curator for the department of paintings, Rijksmuseum Amsterdam; and Professor for technical art history, University of Amsterdam.

Appendix B: CEDR Workshop Program

CEDR Workshop Schedule

Wednesday, May 4, 2016

Participants arrive, check in at the Wyndham Philadelphia Historic District, 400 Arch Street, Philadelphia, PA 19106. Tel. 1-215-923-8660.

Welcome Reception & Dinner

Reception 6:00pm – 7:00pm, Dinner 7:00pm – 8:00pm

Positano Coast Restaurant, 212 Walnut St., Philadelphia, PA 19106.

Thursday, May 5

Chemical Heritage Foundation, Franklin Rooms I & II.

8:00 – 9:00am, Continental breakfast.

9:00 – 10:30am

Welcome – Carsten Reinhardt, President, Chemical Heritage Foundation

"The Vision for a Chymical Encyclopedia, Database, and Repository," James R. Voelkel, Chemical Heritage Foundation & William R. Newman, Indiana University.

10:30 – 11:00am Coffee Break

11:00 – 12:30pm Introduction of Participants

Lawrence Principe, Professor of History of Science and Chemistry, The Johns Hopkins University.

Jennifer Rampling, Assistant Professor of History, Princeton University.

Pamela Smith, Professor of History and Director of the Making and Knowing Project, Columbia University.

Marcos Martinón-Torres, Professor of Archaeological Science, University College London.

Gabriele Ferrario, Research Associate, Taylor-Schechter Genizah Research Unit, Cambridge University Library.

Michelle DiMeo, Curator of Digital Collections, Othmer Library of Chemical History, Chemical Heritage Foundation

12:30 – 1:30pm Lunch

1:30 – 3:00pm Introduction of Participants

Arie Wallert, Curator for the department of paintings, Rijksmuseum Amsterdam; and Professor for technical art history, University of Amsterdam.

Nicolas Thomas, Faculty Member, Institut National de Recherches Archéologiques Préventives, Centre-île de france.

Doris Oltrogge, Cologne Institute of Conservation Sciences, Cologne, Germany.

Marjolijn Bol, Postdoctoral Researcher, Department of Conservation & Restoration, University of Amsterdam.

Jenny Boulboullé, Postdoctoral Scholar, Making and Knowing Project, Columbia University

Donna Bilak, Postdoctoral Scholar, Making and Knowing Project, Columbia University.

3:00 - 3:30pm Coffee Break

3:30 – 5:00pm Introduction of Participants

Wally Hooper, Project Manager/Programmer/Analyst, Chymistry of Isaac Newton Project, Indiana University.

Michelle Dalmau, Head, Digital Collections Services, Indiana University, Bloomington.

Matteo Martelli, Research Associate, Institut für Klassische Philologie, Humboldt-Universität zu Berlin.

Will Cowan, Head of Software Development (Library Technologies), Indiana University Bloomington Libraries.

Joel Klein, Postdoctoral Scholar, Making and Knowing Project, Columbia University.

Seth Rasmussen, Professor of Chemistry, North Dakota State University

Peter J. Ramberg, Professor of History of Science, Truman State University

Dinner 5:30pm

Cuba Libre Restaurant, 10 S. 2nd St., Philadelphia, PA 19106.

Friday, May 6

8:00 – 9:00am Continental breakfast.

9:00 – 10:30am **Encyclopedia.** Editorial structure, workflow, topics.

"The Stanford Encyclopedia of Philosophy," Uri Nodelman, Senior editor, Stanford Encyclopedia of Philosophy.

Potential issues for discussion:

- Will there be articles of a biographical character or just materials, techniques, apparatus?

- What structure do we have for interdisciplinary teamwork in writing the encyclopedia entries? Should we pair chemists (or archeologists or conservators) and historians in some cases? How do we achieve a proper mix of history and scientific rigor – what exactly are we looking for (models of Figala/Priesner, Wietschorek, Wolfgang Schneider, Eklund, SEP).

10:30 – 11:00am Coffee Break

11:00 – 12:00pm **Encyclopedia.** Editorial structure, workflow, topics. (con't)

"Byzantine and Syriac lexica on alchemy: problems and perspectives," Matteo Martelli, Research Associate, Institut für Klassische Philologie, Humboldt-Universität zu Berlin.

12:00 – 1:00pm Lunch

1:00 – 2:30pm **Repository.**

"Text Repositories for Digital Humanities Projects," Michelle Dalmau, Head, Digital Collections Services, Indiana University, Bloomington.

Addendum on non-Indo-European languages, Gabriele Ferrario.

Potential issues for discussion:

1. In the case of edited manuscripts, what do we do with critical apparatus? In the case of MSS. with fluid, contaminated traditions, do we include more than one version? What is a text anyway?
2. Can we use something like the Hathi-trust approach for material under copyright? Non-consumptive analysis.
3. Problem of multiple languages, some of which are not Indo-European. What do we use for headwords? Can non-European documents go into the repository?

2:30 – 3:00pm Coffee Break

3:00 – 5:00pm **Database.**

"Databases as research workspaces: Issues illustrated by the Indiana Dictionary Database project," Wally Hooper, Project Manager/Programmer/Analyst, Chymistry of Isaac Newton Project, Indiana University.

"The Genizah Project," Gabriele Ferrario, Research Associate, Taylor-Schechter Genizah Research Unit, Cambridge University Library.

Potential issues for discussion:

- What do we use for database fields? How do we move from the texts in the repository to the database? How best to mine the repository for the database?

- Problem of linkage and disambiguation. What is the role of metadata? What type of headwords – e.g. materials, apparatus, techniques, people?

- Will the database be public, or just for authors?

Dinner 5:30pm

The Little Lion, 243 Chestnut St., Philadelphia, PA 19106

Saturday, May 7

8:00 – 9:00am Continental breakfast.

9:00 – 10:00am **Replication.**

"White Lead," Lawrence Principe, Professor of History of Science and Chemistry, The Johns Hopkins University.

Potential issues for discussion:

1. What is involved in replications? What sort of standards do we need?
2. How do we facilitate shooting of video clips and still images? We will need some sort of standard format. Should we use something like the EVIA model or another approach?

10:00 – 10:30am Coffee Break

10:30 – 12:30pm **Replication (cont.)**

"Alchemy, archaeology and experiment," Marcos Martinón-Torres, Professor of Archaeological Science, University College London; and Nicolas Thomas, Faculty Member, Institut National de Recherches Archéologiques Préventives, Centre-île de france.

Discussion

"Video, and commenting on replication, the EVIA model," Will Cowan, Head of Software Development (Library Technologies), Indiana University Bloomington Libraries

12:30 – 1:30pm Lunch

1:30 – 3:00pm **Chemist-Historian Collaboration**

Seth Rasmussen, Professor of Chemistry, North Dakota State University

Potential issues for discussion:

3. What structure do we have for interdisciplinary teamwork in writing the encyclopedia entries? Should we pair chemists (or archeologists or conservators) and historians in some cases? How do we achieve a proper mix of history and scientific rigor – what exactly are we looking for (models of Figala/Priesner, Wietschorek, Wolfgang Schneider, Eklund, SEP).

4. How do we get chemists and others who are not in the traditional humanities involved? What is the incentive structure? What sort of protocols do we need to get fruitful work out of bench scientists?

3:00 – 3:30pm Coffee Break

3:30 – 5:00pm **CEDR**

Potential issues for discussion:

1. Chymical dictionaries are basis, but what about recipe literature more generally? What about alchemical practicae? What about the books of secrets tradition? Cookery, perfumery, cosmetics?
2. Is integration possible between existing sites and projects, or future ones, and if so, how?

Dinner 5:30pm

The Gaslight Restaurant, 120 Market St., Philadelphia.

Appendix C: CEDR Prototypes

CEDR Website

Figure 1. CEDR website home page

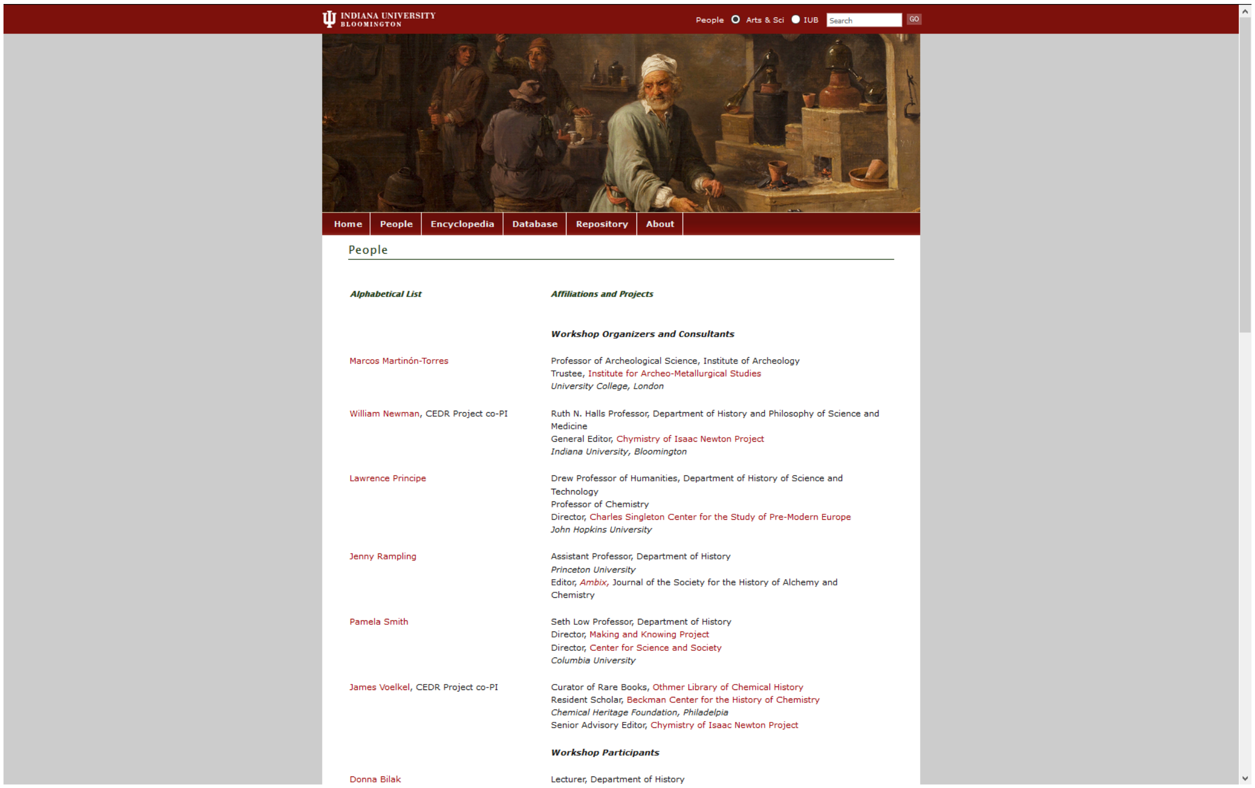
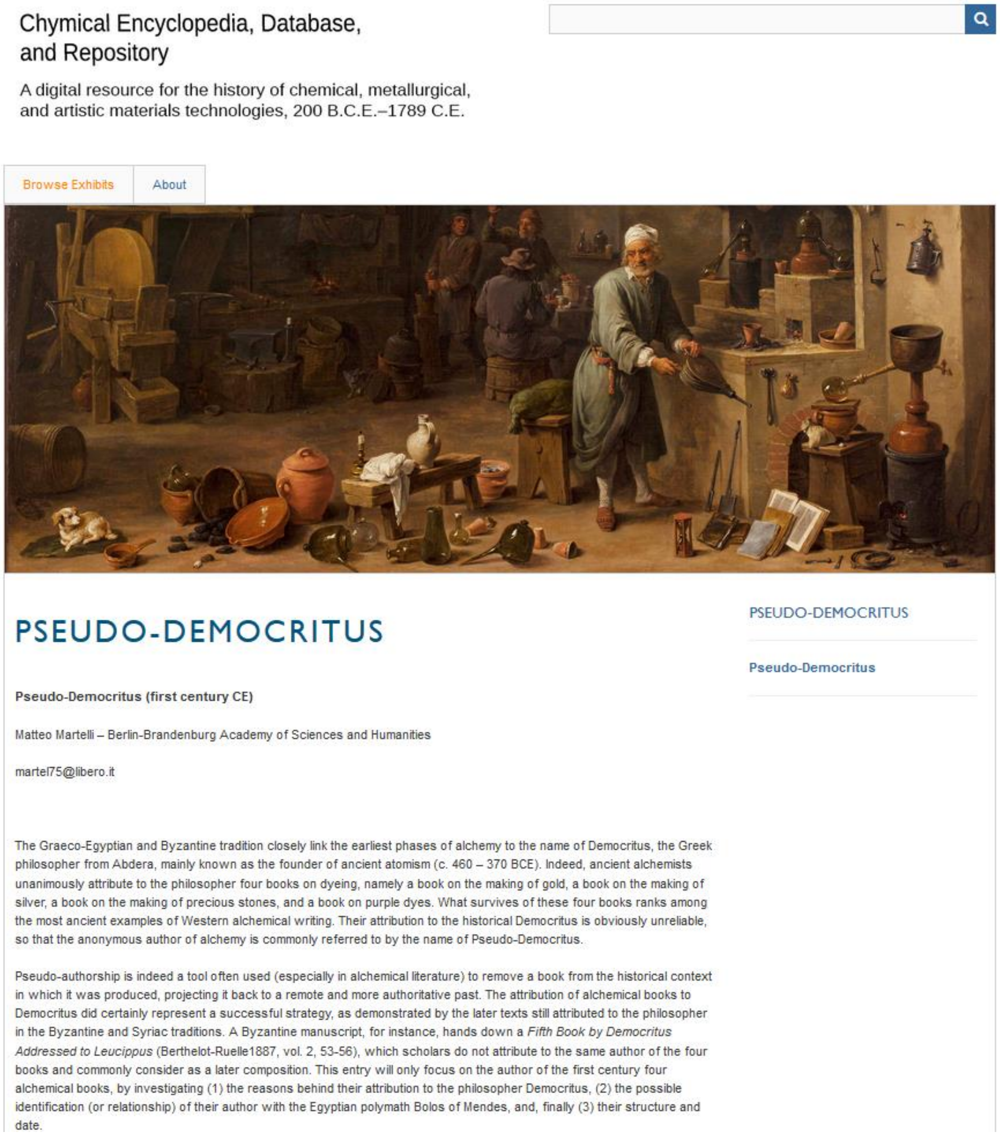


Figure 2. People page from CEDR website

Encyclopedia



Chymical Encyclopedia, Database, and Repository

A digital resource for the history of chemical, metallurgical, and artistic materials technologies, 200 B.C.E.–1789 C.E.

Browse Exhibits | About

PSEUDO-DEMOCRITUS

Pseudo-Democritus (first century CE)

Matteo Martelli – Berlin-Brandenburg Academy of Sciences and Humanities

martel75@libero.it

The Graeco-Egyptian and Byzantine tradition closely link the earliest phases of alchemy to the name of Democritus, the Greek philosopher from Abdera, mainly known as the founder of ancient atomism (c. 460 – 370 BCE). Indeed, ancient alchemists unanimously attribute to the philosopher four books on dyeing, namely a book on the making of gold, a book on the making of silver, a book on the making of precious stones, and a book on purple dyes. What survives of these four books ranks among the most ancient examples of Western alchemical writing. Their attribution to the historical Democritus is obviously unreliable, so that the anonymous author of alchemy is commonly referred to by the name of Pseudo-Democritus.

Pseudo-authorship is indeed a tool often used (especially in alchemical literature) to remove a book from the historical context in which it was produced, projecting it back to a remote and more authoritative past. The attribution of alchemical books to Democritus did certainly represent a successful strategy, as demonstrated by the later texts still attributed to the philosopher in the Byzantine and Syriac traditions. A Byzantine manuscript, for instance, hands down a *Fifth Book by Democritus Addressed to Leucippus* (Berthelot-Ruelle1887, vol. 2, 53-56), which scholars do not attribute to the same author of the four books and commonly consider as a later composition. This entry will only focus on the author of the first century four alchemical books, by investigating (1) the reasons behind their attribution to the philosopher Democritus, (2) the possible identification (or relationship) of their author with the Egyptian polymath Bolos of Mendes, and, finally (3) their structure and date.

PSEUDO-DEMOCRITUS

Pseudo-Democritus

Figure 3. Opening of Pseudo-Democritus article in the encyclopedia.

1. Democritus the Alchemist

The alchemical sources depict Democritus as a Greek philosopher who was taught alchemy by a Persian magus, and mainly operated in Egypt. The alchemist Synesius (fourth century CE), a commentator of Pseudo-Democritus' alchemical writings, explicitly identifies the author of the four books with the philosopher of Abdera, who, after investigating all the natural questions (*ta physika*), was initiated by the magus Ostanes in temple of Memphis with all the Egyptian priests (Berthelot-Ruelle 1887, vol. 2, 56 = Martelli 2013, 122-123). A later Byzantine chronicle (Syncellus, *Chronography* 297,24-28 Mosshammer) includes the name of other alchemists among Ostanes' pupils in Memphis, in particular Maria the Prophetess and the Egyptian Pammenes (see below, section 3).

In his book on gold making, Pseudo-Democritus himself claims to have travelled to Egypt, where he explained how to deal with natural substances. Democritus is said to have been "friend of the Egyptian kings and to have hold a high position among the prophets" (Zosimus, *First Book of the Final Quittance*; see Festugière 1944, vol. 1, 364). In his alchemical works, he probably addressed Egyptian kings and priests (see Berthelot-Ruelle 1887, vol. 2, 427), and he included himself among the 'prophets' (Berthelot-Ruelle 1887, vol. 2, 47 = Martelli 2013, 95). In the late Hellenistic period, the term 'prophet' (*proph_t_s*) – often attributed to early alchemists, such as Maria, Moses or Isis (Mertens 1989) – did not only refer to high-ranking Egyptian priests, but could also indicate any wise figure expert in different (sometimes presumably sacred) areas of knowledge, from magic to botany (Festugière 1949, 380; Boscherini 2007).

Within this Egyptian framework, Persia too plays a pivotal role in the legendary narrative that reinvented Democritus as an alchemist. The philosopher, in fact, is told to have inherited a threefold formula on the powers of nature from the Persian magus Ostanes. This formula summarized Ostanes' alchemical teaching and somehow guided the practice of his pupil. According to a long narrative included in what has been preserved of the four alchemical books (Berthelot-Ruelle 1887, vol. 2, 42-43 = Martelli 2013, 81-85), the master Ostanes would have died before that Pseudo-Democritus accomplished his alchemical training. Despite his attempts to evoke the spirit of the master from the Ades, Pseudo-Democritus was unable to learn the last secrets, until something prodigious happened in an unnamed Egyptian temple (presumably Memphis): a column broke up, and Pseudo-Democritus found inside the threefold aphorism on nature: "Nature delights in nature, nature conquers nature, nature masters nature". A segment of this aphorism closes each recipe of Pseudo-Democritus' books on gold and silver making (see below, section 3).

As clearly emerges from this short summary, atomism is never mentioned in what survives of ancient alchemical writings, so that this philosophical theory does not provide us with a safe background to the attribution of the alchemical books to Democritus. On the contrary, the alchemical tradition reshaped the figure of the atomist philosopher according to different patterns that were quite popular in the late Hellenistic time.

On the one hand, the Egyptian and Persian elements of the alchemical story fit the late Hellenistic image of Democritus: the philosopher was, indeed, represented as a kind of 'globetrotter,' who visited Egypt to learn geometry, went to Babylonia, and travelled as far as India, where he was educated by the famous gymnosophists (see, e.g., Diog. Laert. IX.35). A relation between the atomist and a certain Eastern wisdom had been already suggested by various sources as early as the second to the first century BCE. Hermippus' work *On Magi* probably set this trend, which emphasized the debt of Greek philosophers to foreign wisdom, thus reacting against a certain Heliocentric bias of Greek classical historiography (Bidez-Cumont 1938, vol. 1, 167-168; Ribichini 2001).

On the other hand, Democritus' alchemical skills seem to be part of a wider range of technical competences, which Hellenistic and Roman authors tended to attribute to the atomist. In fact, Democritus started being praised for his expertise in different 'crafts' (*t_chnai*), such as various medical areas (Gemelli Marciano 2007, 213-224), agriculture (Wellmann 1921), military art, and the like. The catalogue of Democritus' writings compiled by the Neo-Pythagorean philosopher Trasyllus (first century BCE – fist century CE; see Tarrant 1993, 85-98) and included at the end of Diogenes Laertius' *Life* of the atomist (Diog. Laert. IX.48) lists eight titles under the heading *t_chnika* (i.e. 'books on technical arts'): four books on medicine (touching upon prognosis, regimen and critical times), a book *On Farming*, a book *On Painting* and two writings on warfare. The authenticity of these treatises – of which we often know just the titles or, in some cases, a few scattered fragments – is still debated among scholars (see, e.g., Lezl 2007, 40-41). Despite this uncertainty, the catalogue is certainly representative of the kinds of expertise Democritus was credited with. For example, as far as medicine is concerned, other sources confirm this picture. The Hippocratic letters depict the philosopher while dissecting animals to discover the origins of madness (*Letter* 17; see

Figure 4. View of the organization element and first section in the Pseudo-Democritus article.

Chymical Encyclopedia, Database, and Repository

A digital resource for the history of chemical, metallurgical, and artistic materials technologies, 200 B.C.E.–1789 C.E.

Browse Exhibits | About

REPLICATION OF ISAAC NEWTON'S REGULUS OF ANTIMONY

Regulus of Antimony made by Reduction with Metals

Because of its interesting semi-metallic properties, many early modern alchemists had a serious interest in antimony (see entry under "Antimony"). However, antimony was typically available only as the mineral stibnite (Sb_2S_3) and thus had to be refined in order to arrive at a "regulus" or button of the metalloid. This was typically performed in one of two ways: either by smelting the stibnite with a metal or with crude or partially calcined wine tartar, often in combination with saltpeter; when refined with tartar, the product was referred to as regulus *per se*. If a metal was chosen as the reducing agent, the choice was usually iron, but other metals were also employed for specific purposes. Here we are concerned only with the reguli made by using one or another metal as a reducing agent. Early modern chymists distinguished between "martial regulus," produced by smelting the stibnite with iron, "jovial regulus" refined with tin, "venereal regulus" made with copper, "saturnine regulus" fabricated with lead, and "lunar regulus" made with silver. A serious question emerges as to whether these different reguli actually contained significant quantities of the respective reducing agent after refining, or whether the initial metal remained only as a trace element. In order to approach this question, successive attempts were made, based on the comprehensive experimental laboratory notebooks left by Isaac Newton (Cambridge University Additional MSS. 3973 and 3975, edited on the *Chymistry of Isaac Newton* site at www.chymistry.org).

Replication

All reguli were made using Newton's recipe in Portsmouth MS Add. 3975, folios 42r and 42v. Unless noted, all samples were made using lab-grade chemicals. Newton suggests using small quantities of metal to produce the best regulus, so we used between 1 and 2 ounces of stibnite for each reaction. All weights are measured in grams. As one of the byproducts of the reaction, antimony trioxide [Sb_2O_3] is a toxic gas, all replications save the last were performed in a small electrical furnace that was placed inside of a fume hood. Newton lists four metals with which reguli may be produced: iron, copper, tin, and lead. In each of these cases, antimony has a higher electron affinity than the reducing metals, so a replacement reaction occurs: the antimony is reduced by the oxidizing metal of choice, giving elemental antimony and a metallic sulfur compound. As will be shown in the results, however, the process does not produce a pure elemental antimony, as large quantities of the reducing metal are still present in the regulus. Part of the antimony is also released as the gaseous antimony trioxide.

REGULUS OF ANTIMONY MADE WITH VARIOUS METALS

Replication of Isaac Newton's Regulus of Antimony

Figure 5. Article on the replication of Isaac Newton's experiments with the regulus of antimony.

Replication

All reguli were made using Newton's recipe in Portsmouth MS Add. 3975, folios 42r and 42v. Unless noted, all samples were made using lab-grade chemicals. Newton suggests using small quantities of metal to produce the best regulus, so we used between 1 and 2 ounces of stibnite for each reaction. All weights are measured in grams. As one of the byproducts of the reaction, antimony trioxide [Sb_2O_3] is a toxic gas, all replications save the last were performed in a small electrical furnace that was placed inside of a fume hood. Newton lists four metals with which reguli may be produced: iron, copper, tin, and lead. In each of these cases, antimony has a higher electron affinity than the reducing metals, so a replacement reaction occurs: the antimony is reduced by the oxidizing metal of choice, giving elemental antimony and a metallic sulfur compound. As will be shown in the results, however, the process does not produce a pure elemental antimony, as large quantities of the reducing metal are still present in the regulus. Part of the antimony is also released as the gaseous antimony trioxide.

Figure 1: Regulus of Antimony made with Copper. Regulus (silver) is visible only after slag (dark grey) was split in two with a wrench.

Regulus of Copper: $Sb_2S_3 + 3\,Cu \rightarrow 3CuS + 2Sb$

Newton gives a ratio of ½ part Copper for every 1 part Stibnite to be fluxed with ¼ part potassium nitrate [KNO_3]. Mixed 15.467 g copper with 31.04 g ground stibnite, then heated in an electric furnace at 800 degrees Celsius for 5 minutes, until the metals were molten. A flux of potassium nitrate—Newton's saltpeter—weighing 7.840 g was added to the crucible. The mixture was then returned to the furnace, and the temperature adjusted to 1,000 degrees Celsius. When the furnace reached 1,000 degrees Celsius (approximately 26 minutes), the crucible was removed and the liquid poured off into a greased crucible.
The receiving crucible split when the molten material was added. Upon breaking, we found a dark grey slag with no apparent regulus. The regulus became apparent only after splitting the product in two. Inside the slag was a bright silver-colored regulus with a crystalline structure (Figure 1).

Figure 2: Regulus of Antimony made with Copper. Slag has been removed, and regulus split to show interior. The interior surface of the regulus has been polished slightly to facilitate examination with SEM

Regulus of Copper: $Sb_2S_3 + 3\,Cu \rightarrow 3CuS + 2Sb$

We attempted to produce a second regulus with copper, this time doubling the weights and using 62.000 g of stibnite, 30.940 g copper, and 15.677 g potassium nitrate. Once again, the stibnite and copper were mixed and placed in a furnace at 800 degrees Celsius for 5 minutes until molten, then fluxed with the potassium nitrate, and the temperature was raised to 1,000 degrees Celsius. As Newton does not specify how long the mixture should remain in the heat, we left it in for 15 additional minutes after the temperature reached 1,000 degrees, for a total of 38 minutes. The molten material was then poured into a greased crucible.

We used a different type of receiving crucible, which did not crack upon contact with the regulus and slag. When the product was cool, we removed it from the crucible and broke off the slag. The slag was dark grey, and inside was a well-formed regulus, approximately the size of a quarter, but slightly thicker (Figure 2). Once the majority of the slag had been removed from the regulus, it was weighed, and found to be approximately 11 g. Newton says that 12 oz. of antimony reduced with copper should produce a regulus of approximately 3 1/3 oz.: approximately 27.5% of the starting weight. We did not fare as well: our regulus was only 17.75% of the starting weight of stibnite. The theoretical yield was 39.5 g Sb and the observed yield was 11.0 g, or 28 %.

Regulus of Tin: $Sb_2S_3 + 3Sn \rightarrow 3SnS + 2Sb$

Newton uses 5 1/3 oz. tin and 3 oz. potassium nitrate for every 12 oz. of stibnite. We combined 30.997 g stibnite with 13.820 g tin and put them in a furnace at 800 degrees for 5 minutes, at which point they were molten. We added the potassium nitrate, returned the crucible to the furnace, and raised the temperature to 1,000 degrees Celsius. As we had received decent results leaving the copper regulus in the furnace for an additional 15 minutes after reaching 1,000 degrees, we decided to leave the tin mixture in for 35 minutes, for a total of 1 hour 9 minutes. The molten mixture was poured into a greased crucible.

Figure 6. Illustrations of experimental products from the encyclopedia article on Newton's experiments with the regulus of antimony.

Repository

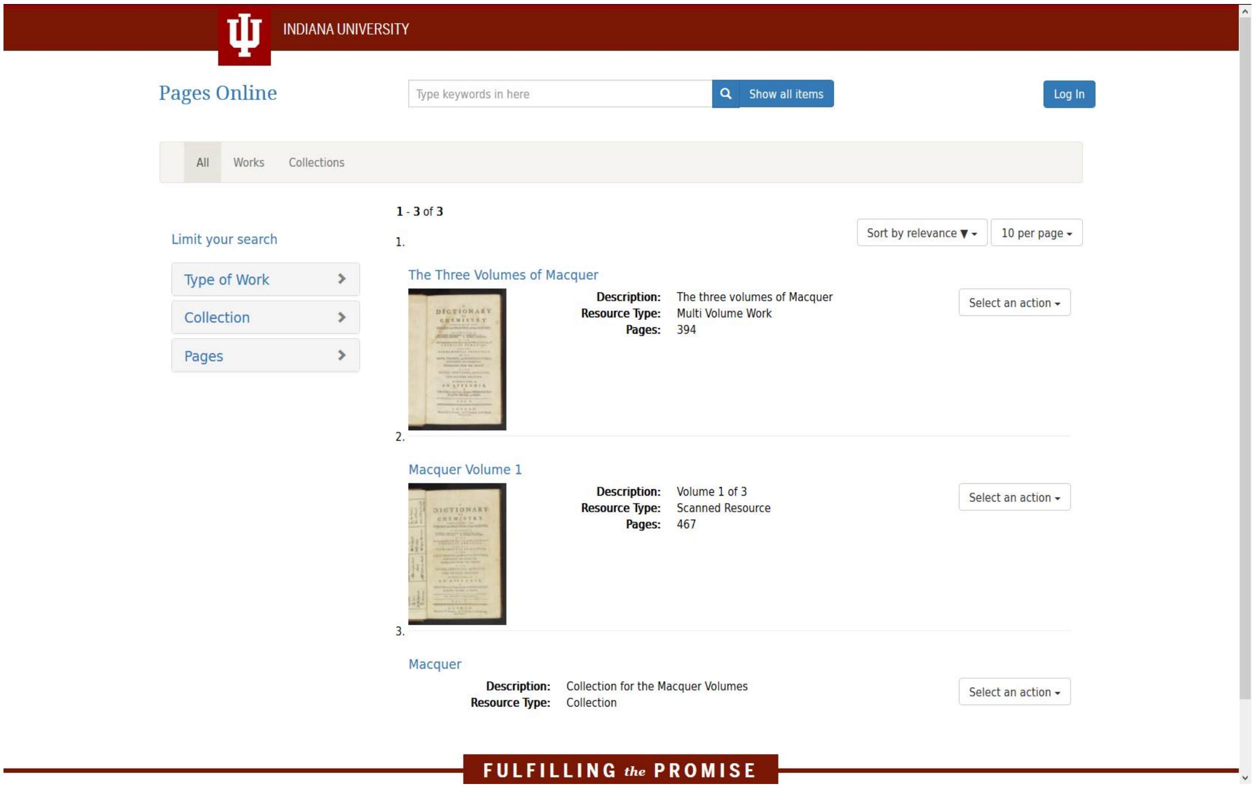


Figure 7. The top-level space in the Pages Repository for page scans of Macquer, Dictionary of Chemistry, *which provides a gateway to 861 pages of material, some of which has been converted to OCR text after a lengthy period of training.*

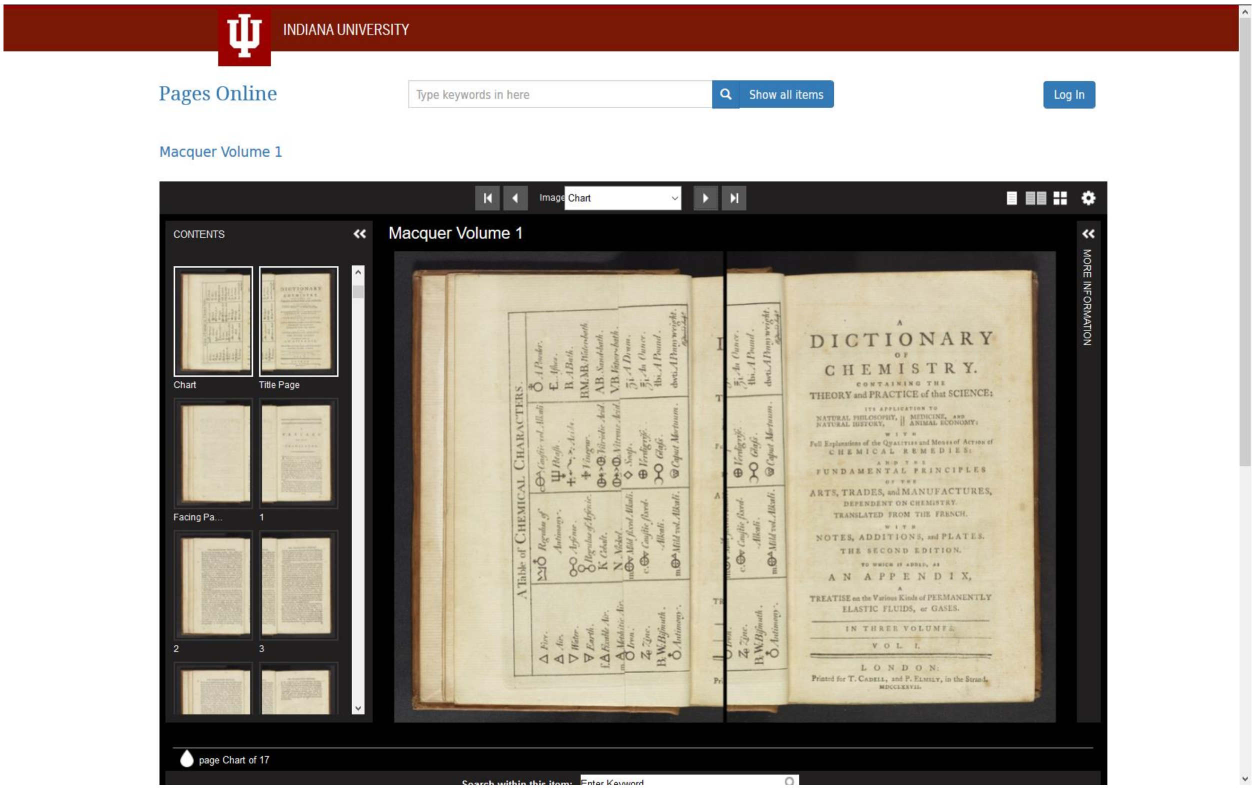


Figure 8. The title page of Macquer's Volume One, with part of a Table of Chemical Characters shown opposite, in the user interface of the Pages Repository.

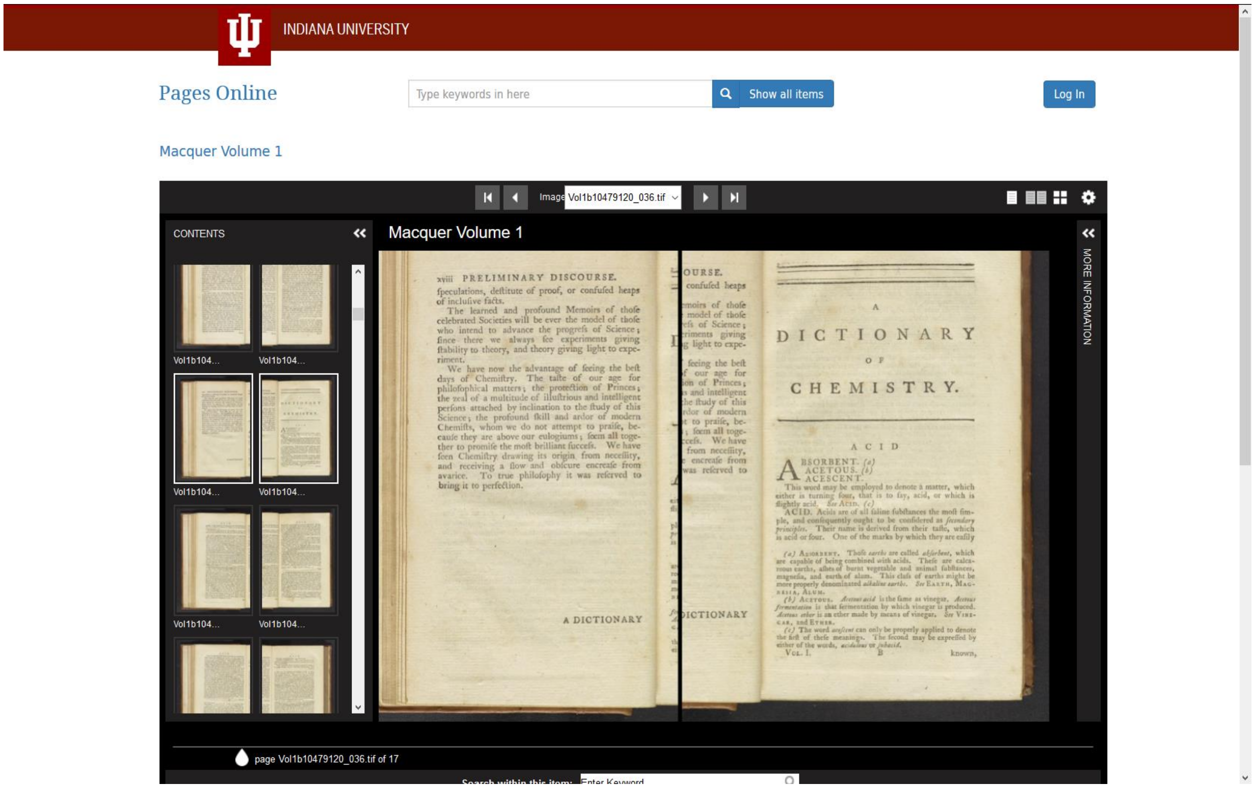


Figure 9. The first entry, ''Acid,'' in the dictionary listing, in the user interface of the Pages Repository.

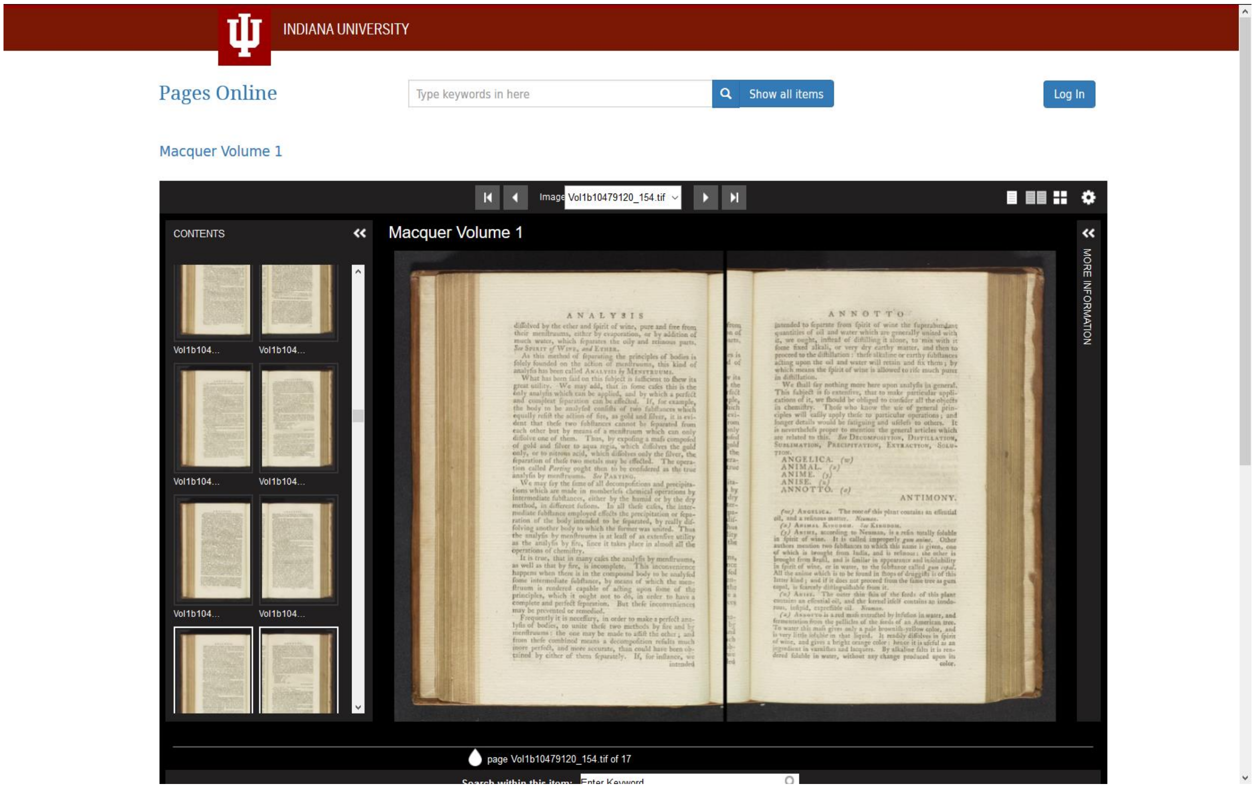
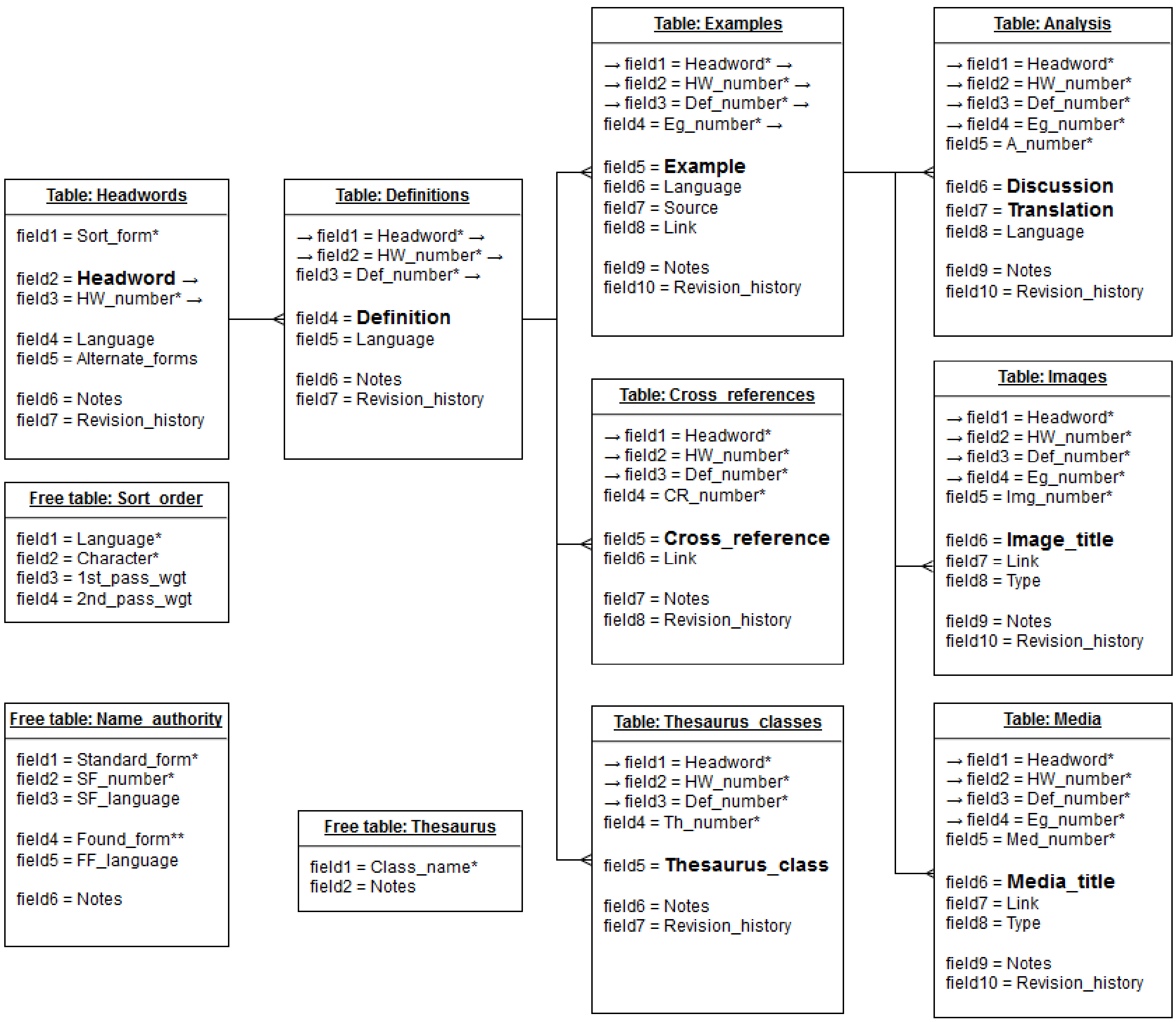


Figure 10. A view in the user interface of the Pages Repository of the articles for "Analysis" and "Antimony," and five unanalyzed headword in between.

Appendix D: Possible Model for a CEDR Relational Database

Basic Design

The purpose of the CEDR database is to clarify the meanings of chymical terms based on analysis of example texts in the original languages, much like a dictionary. In this proposal, the CEDR database is conceived as a relational database outlined in the diagram below. We imagine eight working tables connected in parent-child relationships, and three free tables to organize supporting information. Each table in the diagram lists the fields available in each new record in that table.



The top-level "Headwords" table organizes headwords of interest to CEDR editors, authors, and investigators. This design allows for the possibility that the same lexical form can be analyzed as two different headwords by providing a headword number field, "HW_number" to distinguish them mechanically. Alternate forms of any headword—in the same language or in another—can be recorded beside its accepted primary form. The Sort_form field is generated from the headword using the "Sort_order" table, which is defined by CEDR editors to produce a proper dictionary listing.

Any headword can have more than one sense or definition, based on usage in different authors and contexts. This database design allows any headword to have unlimited definitions stored in separate child records in the "Definitions" table. In turn, each definition record in "Definitions" can have many examples drawn from the literature, or from the archaeological record or replication, which are all stored as separate child records in the "Examples" table. Any definition can have many cross references, and many thesaurus designations, which are stored in the "Cross_references" and "Thesaurus_classes" tables, respectively.

The "Analysis" table, which is a child table of "Examples," allows CEDR investigators, authors, and editors to provide translations of each example in many different languages, and to add unlimited amounts of discussion in separate records, as needed for the purposes of clarification or debate as collaborative work proceeds. The "Examples" table is also the parent table for "Images" (TIFF, PNG, JPG, etc.) and "Media" (sound and video) records, with linking information. The design assumes that the CEDR database user interface will display the images and play the media.

All the definitions, examples, and other information in the database are stored in these eight tables. The relationships between headwords and their child records are managed and distinguished from one another by the presence of fields inherited from parent records. Inherited fields are used to index and sort the child tables and to maintain connections child records and their parents. (Inheritance is indicated in the diagram by arrows; fields used to index a table are marked by an asterisk.)

This design supports work in any language by using language ISO designations and Unicode. Tables provide a Language field where an ISO number can be stored to designate the language being used on that record; this facility could aid in the selection of records and the composition of printed reports.

For editorial work and revision tracking, all tables provide fields for Notes and Revision_history.

This design only addresses possible tables, fields, and indexes. It assumes that the database platform chosen for the final implementation will provide intuitive user interfaces and sophisticated search and reporting functions, which are now prevalent in the relational database world.

Ingestion of Computational Results

We assume that computational harvesting will begin with a selection of target headwords that have counterparts in the "Headword" table. The harvesting algorithm will look for the targets in documents in the CEDR repository or elsewhere, and extract nearby text when found, so the harvested results will probably consist of a series of items that pair a headword with an extracted passage.

The CEDR ingestion program will look at each paired result and locate the headword in "Headwords." The program will add a new, empty, child record in "Definitions" to serve as a placeholder and parent record for a new child record in "Examples," where the harvested passage will be inserted, along with information for the Language, Source, and Link fields.

Human editors will have to complete the ingestion process by reconciling each newly harvested "Examples" record with any existing "Definitions" and "Examples" records found under its parent "Headword" record, with the idea of creating a consistent dictionary entry.